

Datenvisualisierung mit JMP

Patrick René Warnat
HMS Analytical Software GmbH
Rohrbacherstr. 26
Heidelberg
patrick.warnat@analytical-software.de

Zusammenfassung

Das JMP Paket ist ein Softwareprodukt der Firma SAS für Datenvisualisierung und -analyse.

In diesem Beitrag werden die Werkzeuge und die Möglichkeiten zur Datenvisualisierung mit JMP anhand von Beispielen vorgestellt.

Eine Besonderheit von JMP ist die graphische Benutzeroberfläche, die leicht zu bedienen ist und insbesondere auch die interaktive Arbeit an Graphiken unterstützt. So sind Datentabellen und daraus abgeleitete Datenvisualisierungen dynamisch miteinander verbunden: Wird zum Beispiel in einer Datentabelle eine Zeile mit den Daten einer Beobachtung markiert, so wird in einem von der Tabelle zuvor abgeleiteten Streudiagramm gleichzeitig die graphische Repräsentation der Beobachtung hervorgehoben.

Weiterhin kann das Erscheinungsbild einer erstellten Graphik durch eine Reihe von Werkzeugen sehr einfach modifiziert werden, so lassen sich zum Beispiel in Streudiagrammen der auf einer Achse dargestellte Wertebereich durch „Ziehen“ mit der Maus direkt in der Graphik ändern.

Für grundlegende Diagrammtypen (Streudiagramme, Histogramme, Box Plots, ..) ist die Graphikerstellung in der aktuellen Version 8 von JMP mit dem interaktiven „Graph Builder“ noch einfacher geworden.

Über einfache Diagrammtypen hinaus bietet JMP noch einiges mehr. In diesem Beitrag wird dies anhand eines Streifzugs durch unterschiedliche Beispiele gezeigt, unter anderem die Erstellung von animierten Streudiagrammen oder der Export von Bildern oder Animationen zur Verwendung in anderen Applikationen.

Schlüsselwörter: JMP, Graphiken, Datenvisualisierung

1 Einleitung

Datenvisualisierung ist ein mächtiges Werkzeug zur Analyse quantifizierbarer Informationen, da dabei eine besondere Stärke des Menschen genutzt wird: die visuelle Wahrnehmung. Durch die grafische Darstellung von Daten lassen sich Muster, Trends und Zusammenhänge oft schneller und einfacher identifizieren als ohne Visualisierung. Insbesondere zu Beginn einer Datenanalyse ist die Datenvisualisierung eine sehr gute Methode zur explorativen Erkundung von Daten und, darauf aufbauend, zur Hypothesengenerierung. Weiterhin ist die Datenvisualisierung ein unabdingbares Hilfsmittel zur Kommunikation von Ergebnissen einer Datenanalyse in organisations-

internen Besprechungen und Berichten als auch öffentlichen Konferenzen oder Publikationen.

Software-Werkzeuge zur Visualisierung von Daten sollten flexible Möglichkeiten als auch eine einfache Bedienbarkeit bieten. Eine solche Software ist die JMP[®]-Applikation zur Datenanalyse und Visualisierung.

Die Software JMP der Firma SAS bietet umfangreiche Möglichkeiten zur Visualisierung von Daten. Komfortable Visualisierungsmethoden unterstützen die Auswertung von Untersuchungsergebnissen.

Ein besonderes Merkmal der Software ist dabei die leicht zugängliche Bedienung über eine ausgereifte graphische Benutzeroberfläche.

Detaillierte Informationen zu den Möglichkeiten der JMP Software zur Graphikerstellung und Datenvisualisierung finden sich in [1].

Im Folgenden wird eine einführende Übersicht zu ausgewählten Aspekten der Datenvisualisierung mit JMP anhand von Beispielen gegeben.

2 Überblick zur Datenvisualisierung mit JMP

2.1 Erstellung Graphiken mit dem Graph-Builder

Mit der Version 8 der JMP – Software ist der Einstieg in die Graphik-Erstellung mit der JMP-Software noch einfacher geworden: Der sogenannte „Graph-Builder“ erlaubt die Erstellung von Diagrammen mit Hilfe eines interaktiven Dialogs. Verschiedene Diagramm-Elemente sind im Graph Builder verfügbar („scatterplots“, „box plots“, „histograms“, „bar charts“, „lines“), um univariate und multivariate Sachverhalte und Zusammenhänge darzustellen. Daten können dabei gruppiert visualisiert werden, wobei sowohl die überlagerte als auch die getrennte Darstellung von Diagrammen möglich ist. Bei Aufruf des Graph Builders über den Menüpunkt „Graph/Graphik erstellen“ erscheint für die aktuell geöffnete Datentabelle ein Dialog mit einer interaktiven Diagrammvorschau, wie in Abbildung 1 gezeigt.

Durch Auswahl von Spalten der aktuell geöffneten Datentabelle sowie Ziehen und Loslassen von Spalten in bestimmte Bereiche der Diagrammvorschau lässt sich eine Graphik interaktiv erstellen. Die Eigenschaft der einzelnen Diagrammelemente (z. B. Farben) lassen sich über Kontextmenüs steuern.

Ein Beispiel für ein mit dem „Graph Builder“ erzeugtes Diagramm zeigt Abbildung 2.

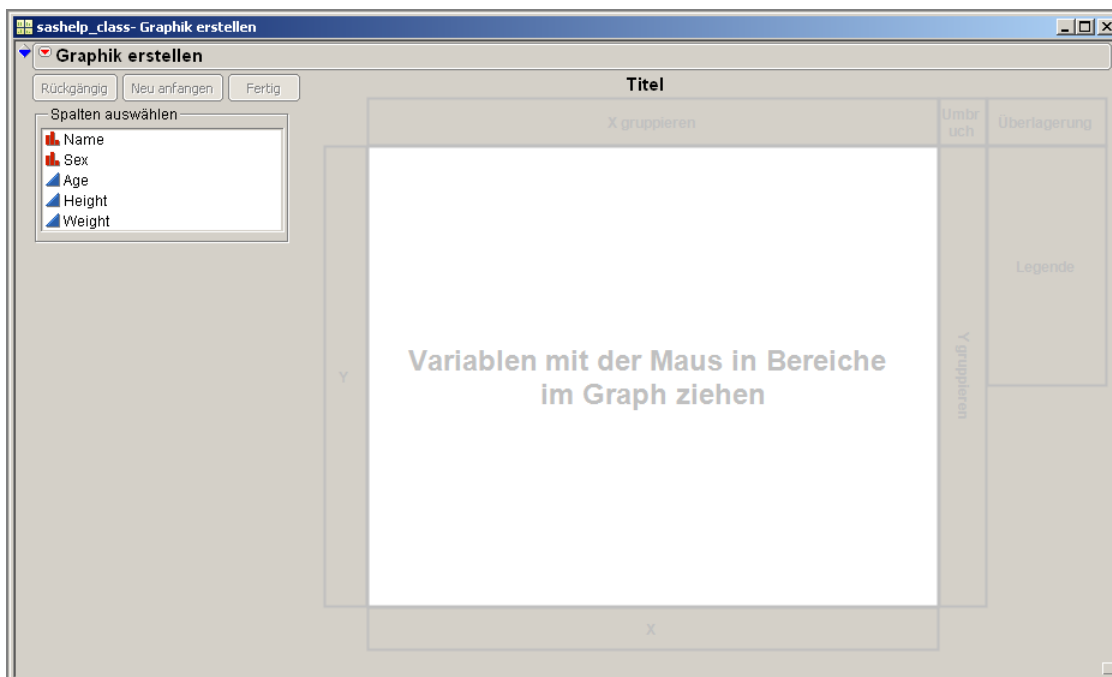


Abbildung 1: Der „Graph Builder“-Dialog zur interaktiven Diagrammerstellung

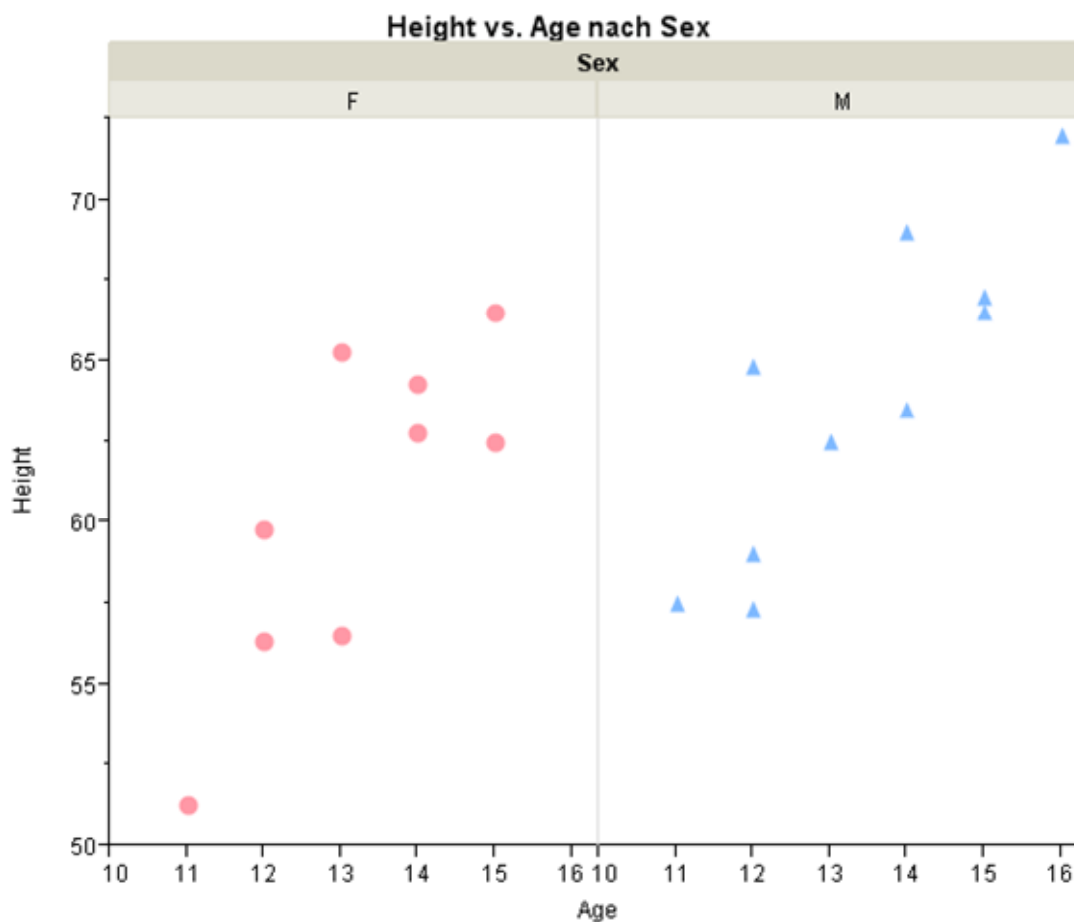


Abbildung 2: Beispiel für ein mit dem „Graph Builder“ erzeugtes Diagramm

2.2 Interaktive Verknüpfung Daten und Graphiken

Eine Besonderheit von JMP ist die graphische Benutzeroberfläche, die leicht zu bedienen ist und insbesondere auch die interaktive Arbeit an Graphiken unterstützt. So sind Datentabellen und daraus abgeleitete Datenvisualisierungen dynamisch miteinander verbunden: Wird zum Beispiel in einer Datentabelle eine Zeile mit den Daten einer Beobachtung markiert, so wird in einem von der Tabelle zuvor abgeleiteten Streudiagramm gleichzeitig die graphische Repräsentation der Beobachtung hervorgehoben. Auch umgekehrt funktioniert diese interaktive Verknüpfung: Elemente in einer Graphik werden mit der Maus angewählt, und gleichzeitig werden die entsprechenden Beobachtungen in der Datentabelle hervorgehoben, wie in Abbildung 3 gezeigt.

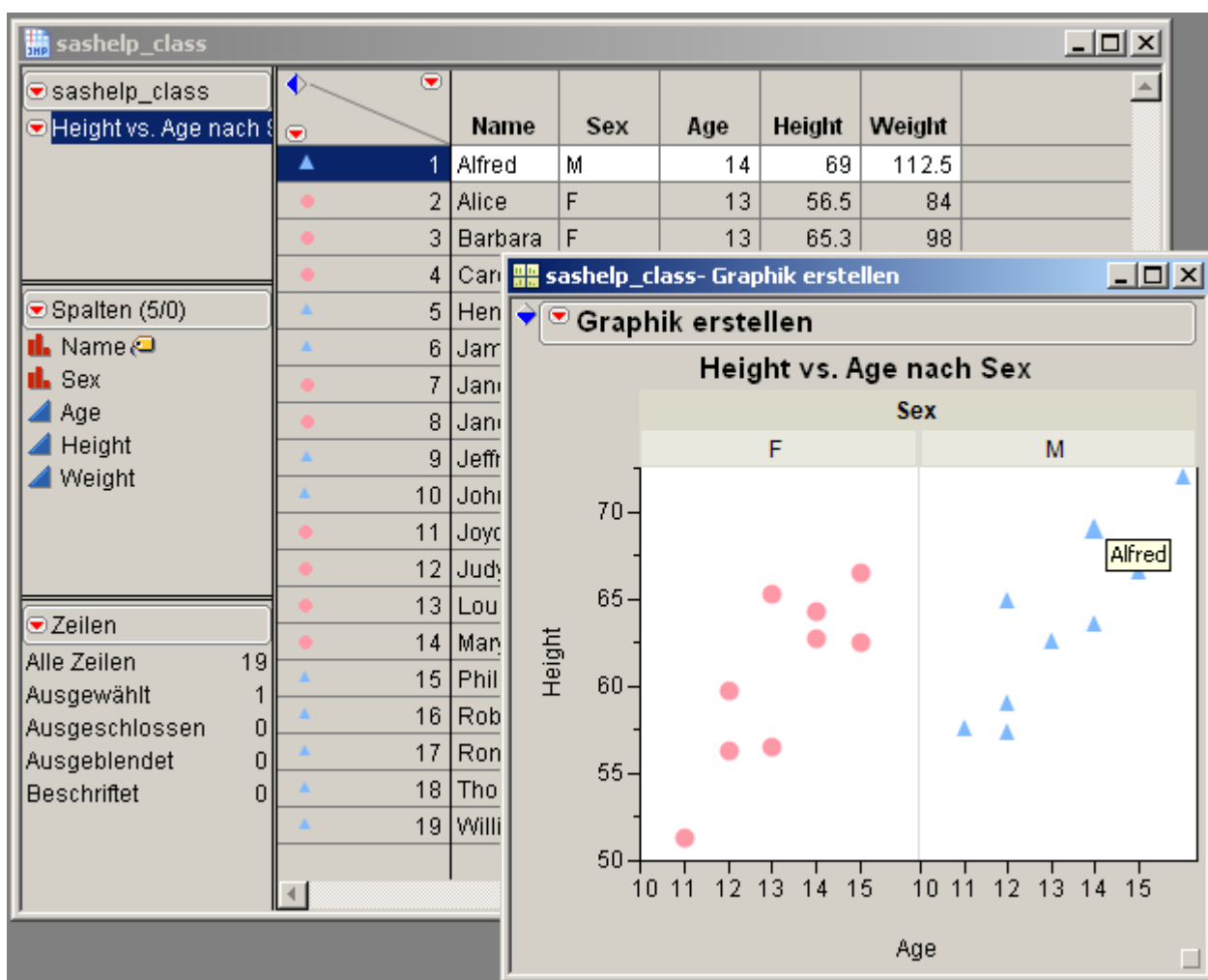


Abbildung 3: Beispiel für die interaktive Verknüpfung von Graphiken und Beobachtungen in einer Datentabelle: Elemente in einer Graphik werden mit der Maus angewählt, und gleichzeitig werden die entsprechenden Beobachtungen in der Datentabelle hervorgehoben.

2.3 Interaktive Änderung von Graphiken mit allgemeinen Graphik – Werkzeugen

Jede in JMP erstellte Graphik kann innerhalb von JMP nachträglich bearbeitet werden. Neben Diagrammtyp-spezifischen Kontextmenüs bietet JMP mehrere Werkzeuge, die die direkte Manipulation einer Graphik mit der Maus ermöglichen.

So kann mit dem „Hand“-Werkzeug etwa der dargestellte Achsenabschnitt in einem Diagramm beeinflusst werden oder mit einer Lupe kann in eine Graphik „hinein-gezoomt“ werden.

Weitere Möglichkeiten bestehen durch so genannte Annotations-Werkzeuge. Damit lassen sich Linien, Pfeile oder beliebige Texte direkt in eine Graphik setzen.

2.5 Speicherung von Graphiken und Graphik-Export

Eine erstellte Graphik lässt sich entweder als in einem JMP-spezifischen Format („JMP-Report“) speichern oder in einem der gängigen Graphik-Formate wie jpg, png oder gif. Die Speicherung als JMP-Report hat den Vorteil, dass die interaktive Bearbeitung in JMP auch bei einer späteren JMP-Session weiterhin möglich ist. Mit der Speicherung in ein externes Format geht die Möglichkeit der interaktiven Bearbeitung verloren, allerdings kann eine exportierte Graphik dafür in jeder anderen Applikationen gelesen werden, die das ausgewählte Graphik-Format unterstützt.

2.6 Skript-Speicherung und Skripting

Graphiken können auch als so genanntes Skript gespeichert werden. Dabei wird keine Binär-Datei gespeichert, sondern eine Text-Datei, die den Programmcode enthält um die Graphik aus der verwendeten Datentabelle erneut zu erzeugen. Der Programmcode ist in der so genannten „JMP Scripting Language“ (JSL) implementiert, eine Art Steuersprache für JMP. JSL-Skripts können, ähnlich zu Visual Basic-Makros in Microsoft-Office, verwendet werden um Vorgänge in JMP zu automatisieren oder auch eigene Abläufe und Auswertungen zu entwickeln.

Wenn man das JSL-Skript einer Graphik speichert, ist es ein Leichtes eine Graphik erneut aus einer Tabelle zu erstellen, wenn z. B. neue Beobachtungen in die Tabelle eingetragen wurden, ohne die Graphik erneut aufzubauen zu müssen. Dies kann bei komplexen Graphiken einiges an Zeitersparnis erbringen.

2.7 Beispiele Spezial-Graphiken

JMP bietet neben allen gängigen Diagrammtypen einige Spezial-Graphiken. So ist es z. B. mit dem Menübefehl „Graph/Streudiagramm-Matrix“ ein Leichtes eine Matrix mit Streudiagrammen zu erstellen, wie in Abbildung 4 dargestellt.

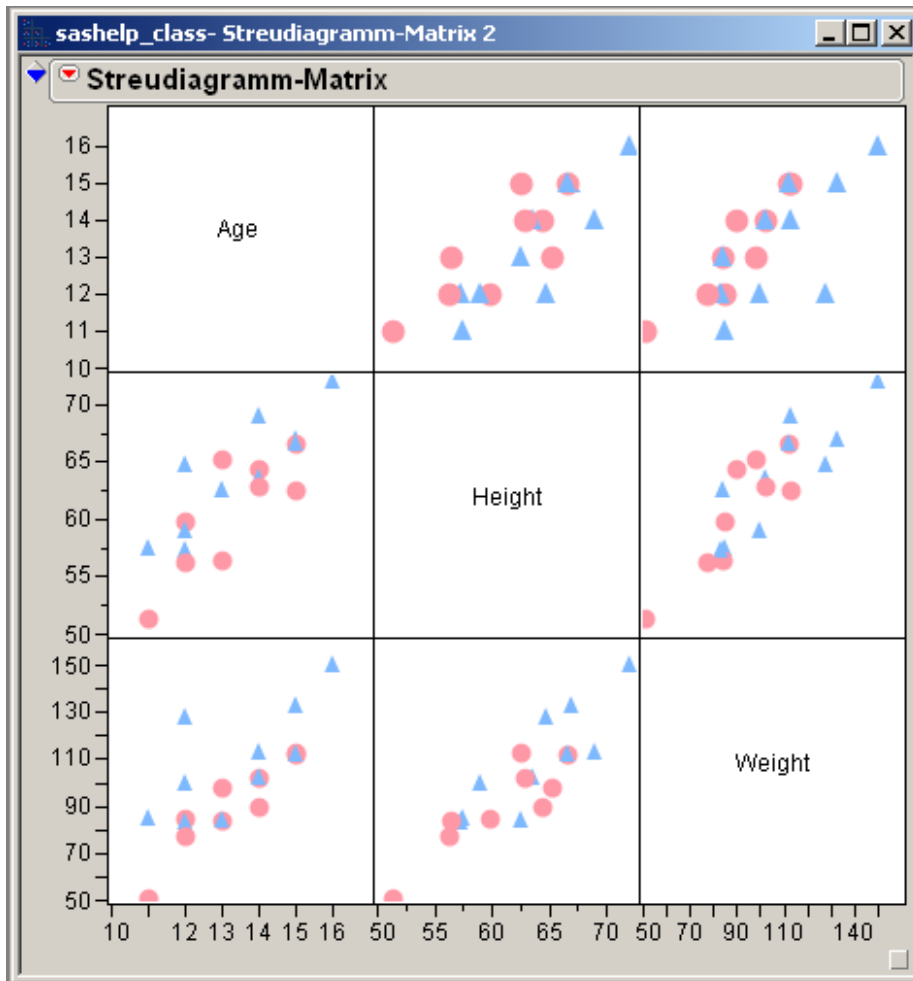


Abbildung 4: Beispiel für eine Streudiagramm-Matrix

Ein weiteres Beispiel für eine Spezial-Graphik ist ein 3D – Streudiagramm wie in Abbildung 5 gezeigt („Graph/3D-Streudiagramm“). Die Lage des „Würfels“ lässt sich interaktiv mit der Maus ändern.

Als letztes Beispiel für eine Spezial-Graphik sei das so genannte „Blasendiagramm“ genannt, welches in einem Streudiagramm für jede Beobachtung eine weitere Variable visualisiert, in dem die Größe der gezeichneten Markierung („Blasen“) variiert wird (Abbildung 6). Als Besonderheit lässt sich dieses Diagramm auch animieren, um einen Verlauf über die Zeit zu visualisieren.

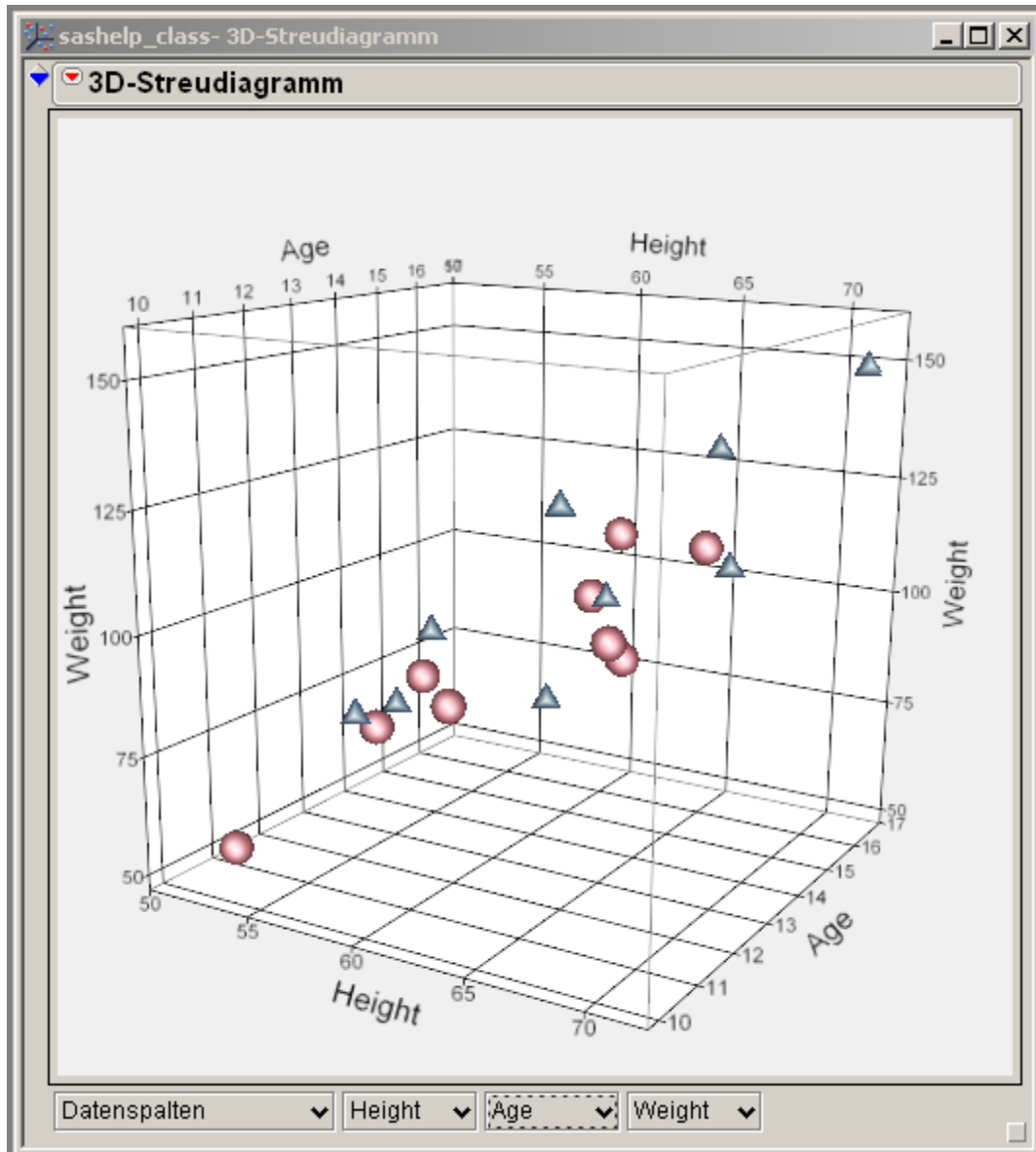


Abbildung 5: Beispiel für ein 3D-Streudiagramm

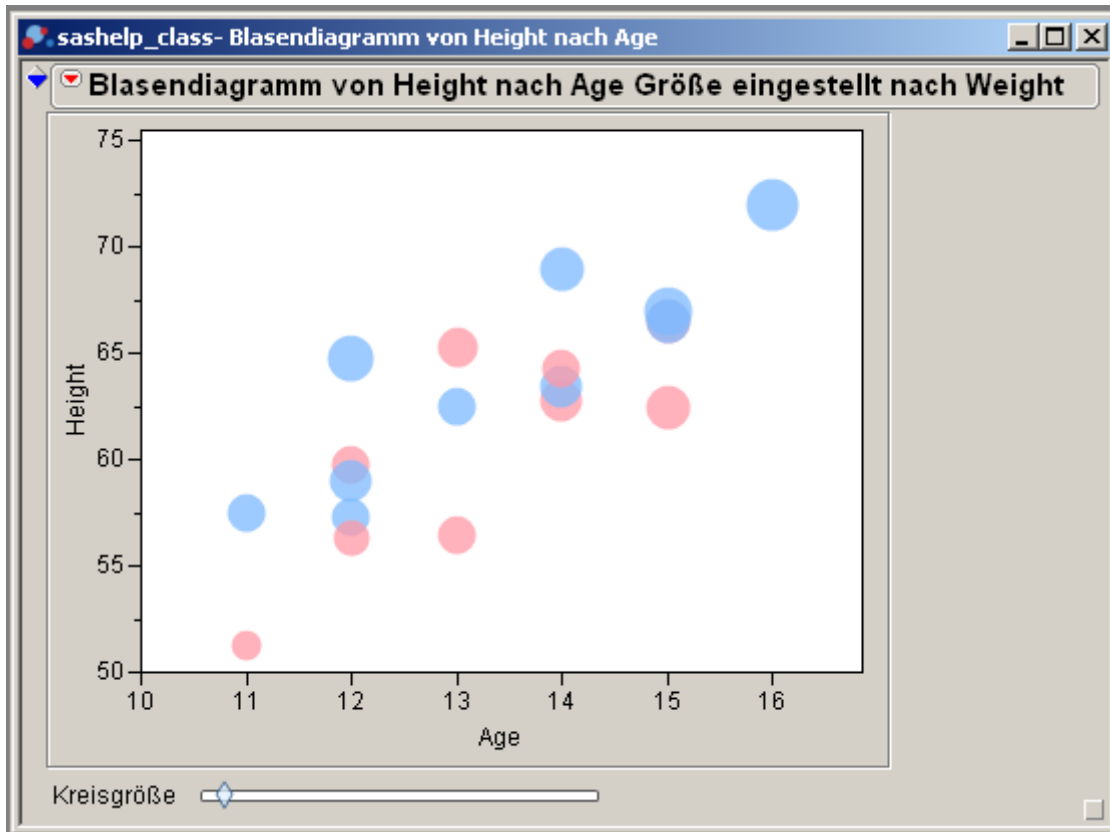


Abbildung 6: Beispiel für ein Blasendiagramm

3 Fazit

Die JMP-Software ist insbesondere für interaktive Datenvisualisierung sehr geeignet, da sie dies mit verschiedenen Werkzeugen sehr gut unterstützt, und dabei noch schnell und einfach zu bedienen ist.

Literatur

- [1] JMP Statistics and Graphics Guide, Release 8. SAS Institute Inc., Cary, NC, USA, 2008.